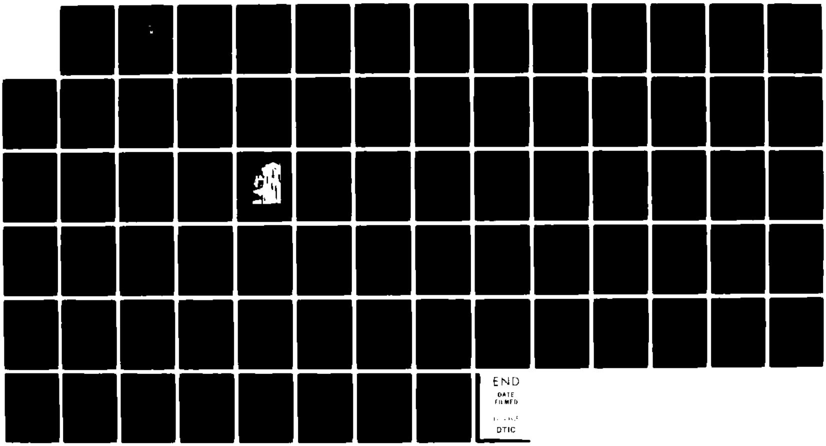
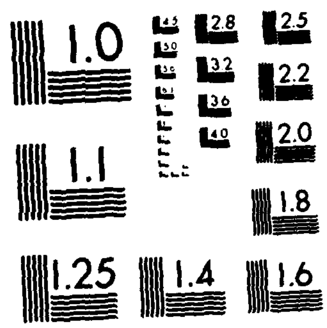
AD-A128 559 SOME EFFECTS OF STRESS ON USERS OF A VOICE RECOGNITION SYSTEM: A PRELIMINARY INQUIRY(U) NAVAL POSTGRADUATE SCHOOL MONTEREY CA B A FRENCH MAR 83 1/1

UNCLASSIFIED F/G 5/10 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

2

NAVAL POSTGRADUATE SCHOOL

Monterey, California

THESIS

SOME EFFECTS OF STRESS ON USERS OF
A VOICE RECOGNITION SYSTEM:
A PRELIMINARY INQUIRY

by

Brian Arthur French

March 1983

Thesis Advisor: G.K. Poock

DISCLAIMER NOTICE

THIS DOCUMENT IS BEST QUALITY PRACTICABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO. AD-A128559	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Some Effects of Stress on Users of a Voice Recognition System: a Preliminary Inquiry		5. TYPE OF REPORT & PERIOD COVERED Master's Thesis; March 1983
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Brian Arthur French		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		12. REPORT DATE March 1983
		13. NUMBER OF PAGES 74
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Voice Data Entry
Speech Recognition
Voice Recognition
Automatic Speech Recognition
Human Factors
Time-induced Stress
Stress

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Recent work with Automatic Speech Recognition has focused on applications and productivity considerations in the man-machine interface. This thesis is an attempt to see if placing users of such equipment under time-induced stress has an effect on their percent correct recognition rates. Subjects were given a message-handling task of fixed length and allowed progressively shorter times to attempt to complete it.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. (continued)

Questionnaire responses indicate stress levels increased with decreased time-allowance; recognition rates decreased as time was reduced.

A

Approved for public release; distribution unlimited.

Some Effects of Stress on Users of a Voice Recognition System: A Preliminary Inquiry

by

Brian Arthur French
Lieutenant, United States Navy
B.A., Yale University, 1975

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN SYSTEMS TECHNOLOGY
(Command, Control and Communications)

from the

NAVAL POSTGRADUATE SCHOOL
March 1983

Author ______________________________

Approved by: ______________________________
Thesis Advisor

____
Second Reader

____
Chairman, Command, Control and Communications
Academic Group

____
Academic Dean

ABSTRACT

Recent work with Automatic Speech Recognition has focused on applications and productivity considerations in the man-machine interface. This thesis is an attempt to see if placing users of such equipment under time-induced stress has an effect on their percentage correct recognition rates. Subjects were given a message-handling task of fixed length and allowed progressively shorter times to attempt to complete it. Questionnaire responses indicate stress levels increased with decreased time-allowance; recognition rates decreased as time was reduced.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

I. INTRODUCTION

Stress has long been an intensely active area of investigation. More recently, the intersection of physiology and psychology has concerned itself with the influence of cognitive and emotive factors on quantifiable variable n the human organism's equation. This research has turned m the most fundamental indices of human reaction (car c rate, electrogalvanic response) to more subtle measures a as various features of the human voice.

Stress, however defined, has clear and important implications for the conduct of military operations. In the warfare environment, one encounters some of the few factors universally acknowledged to cause psychological stress: it is one of the few areas where one knows apriori that life-threatening situations are consistently found. Thus the psycho-physiological correlates of military action have furnished both an area of inquiry and the raw data for those interested in the study of the human under stress. Swedish researchers took an early lead in the investigation of stress in the military environment. [Ref.1]

The investigation of any phenomenon requires first and foremost a statement delimiting the area of research. Stress has yet to receive a widely accepted definition. Stress

research tends to divide into two camps: those interested in physiological stress and those pursuing the psychological variety. Both claim the territority of psychophysiology. Beyond this, the definition of stress breaks down, in the main, into an examination of either a stimulus or a response: there appears to be no common ground. One thing which emerges clearly from published research in the area is that a synthetic conception of stress is in order, one broad enough to cover the wide (and widening) area of inquiry already embarked upon, and which would permit breakdown into narrower, more specialized areas of interest.

Given the vagaries of the published resources currently available, it is necessary within the context of the present paper to work within an operationally-d fined, problem-oriented view of stress. From a military commander's point of view, an area of great potential interest is the interaction of stress with voice-input to computer systems. As miniaturization and expanding technology have made computers a more widespread phenomenon in the military environment, the limiting factor (in the military just as with civilian applications) continues to be the human interface with the ever- expanding base of information and decision aids available. Voice-input to computers provides the field user with a higher-order interface, at a level above even user-friendly query languages. This type of interface has the potential to facilitate the user's ability

to access and manipulate the vast resources now becoming available to him: it is the next logical step in the evolution of man-machine compatability. Computers are now able to handle problems which not long ago would have absorbed an enormous amount of man-hours, but the man is still "in the loop".

> ...though we hail the advent of the postindustrial society, in which substantially all our essential human needs will be produced without a great amount of human effort, the fact remains that the millenium is not here yet. Human beings are not yet obsolescent as important contributors to the productive process. The quality of performance by men and women, whether in manufacturing or servicing, still accounts for a very large part in the variation in productivity, however one chooses to define the criteria of productivity - and this condition will continue to prevail far into the future. [Ref. 2: p. XIII]

Automatic Speech Recognition (ASR) is a demonstrated, on-line technology of clear interest to the military commander. (For a discussion of the range of voice input applications, the reader is referred to References [3] and [4]) The rapidity with which current data processing and computer networks now operate necessarily imply that the machinery spends an inordinate part of its time waiting for input from the human operator. Speeding up this aspect of the interface, a quantum leap in user-accomodation, is one of the major contributions to be expected (and even now available) from speech recognizers linking the human to the computer. Speech recogniton systems, while they generally do not affect the internal workings of the computer, present

their greatest potential advantage in increasing the efficiency of the total human/machine system. Industry was quick to recognize this potential: voice input systems are presently in use in various sorting and assembly-line applications.

Computer input by voice presents several advantages in the military context, especially in the Command Post environment. It frees operators' hands to enable other concurrent tasks, such as transcription or manipulation of other equipment, increasing productivity in both areas.[Ref. 3] Just as importantly, it provides the potential for the voice-operator to maintain a graphic or figurative plot of a situation he may be monitoring at the terminal. This would allow others not familiar or comfortable with the computer to comprehend and follow what the operator may see unfolding at his work station, without requiring lengthy and distracting explanations from the operator. Voice-input could enable the operator to continue his task at the terminal and simultaneously manipulate a visual representation of the problem he is involved in, for others' benefit. This is a potential boon in the period of transition from a symbolic gestalt to an era of much more widespread computer literacy.

Moreover, cognitive psychology has presented a scientific model for task suitability to different modalities which takes the criteria for task assignment far

beyond the obvious basis of "ease". Sandry and Wickens [Ref. 5: pp.8-10] have proposed the concept of "ideomotor compatability": human input/output modes may be matched to the senses and hemispheres of the brain which most readily process them. Quicker, more accurate responses result when responses are matched to stimuli (such as hearing and speaking). Short term memory for verbal information is significantly improved when that information is presented auditorily.

There are obvious correlates of ideomotor compatability in a situation such as a Command Post. The commander gives an operator verbal instructions regarding a task, and this operator then executes these with voice input to a terminal. In competing for the human operator's mental resources, it is Sandry and Wickens' thesis that no time-sharing decrement results when competing tasks are "ideomotor compatible".

The automated resources available to the commander in the current environment are expanding and gaining increasingly widespread acceptance. The applicability of voice input to computer systems is therefore both intuitively and logically attractive, and has been gaining increasing attention. Stress, however defined, is universally conceded to be a factor in the human side of this equation, and therefore to has implications of military interest in the command and control environment.

Given these considerations, what then might be the effect of stress on an operator of a voice input system? In various areas, research on psychological stress abounds, and researchers are now turning their attentions specifically to the effects of stress on the voice. The published research in this latter area is, in the main, inconclusive. It is the intent of this thesis to see if some relationship exists between stress incident on a voice-input operator and his success rate with an automatic speech recognizer. It is to be emphasized that virtually no work has been published in this particular area. The current study proposes itself as a preliminary inquiry into a phenomenon of potential significance, and one which has as yet received little experimental attention.

II. VOICE RECOGNITION AND STRESS

A. DEFINITION

Central to any study of the interaction of stress and the output of voice recognizers is a definition of what is to be understood by "stress". As stated (cf. supra), the disagreement in defining stress is so fundamental that even recent treatments differ in classing stress as stimulus or response. In certain areas, however, consensus emerges.

Stress in the physiological sense is more clear-cut than the psychological. Physiologists may refer unambiguously to human stressors such as heat, pressure, electric shock, and the like, in the way of stimuli. The resultant reactions are also referred to as stress, and are measured in cardio-pulmonary parameters, pupil dilation, and similar observable and quantifiable phenomena. When speaking of physiological stress, it suffices to specify whether stimulus or response is meant. Many stressor-agents have been extensively studied both in this country and abroad for their military impacts.

The grey area is entered with psychological stress and its causes or manifestations. Various researchers define stress as the organism's reaction (coping or adaptive) to an unfriendly environment. Others define stress as the inimical influence of the environment on the individual. In recent usage, the former definition appears to have gained

ascendancy over the latter, although the debate is far from settled.

It is therefore incumbent to define psychological "stress" before proceeding further. The definiton must be sufficiently broad to span a wide gamut of potential phenomena, and thus preserve meaingfulness. In addition, it must be sufficiently narrow to be useful and workable within the necessarily limited scope of this experiment. A synthetic definition of stress will here mean "the arousal produced within the individual by unwelcome stimulus or threat". This is the operational definition which will be adopted in an attempt to examine the end-effect of producing this type of arousal in subjects interacting with an automatic speech recognizer.

There is extensive literature on the subject of psychological stress (much of which falls witnin the definition above) and its correlates, especially the physiological correlates. The Navy has done extensive work in the field, most notably in San Diego. [Ref. 6] Psychological stress may be further broken down into interesting subcategories. 1) Situational stress is generally held to be the influence on the individual of unfavorable environmental (but non-physical) factors. These factors are beyond the individual's control and are generally thought to consist of a certain set of circumstances such as public speaking, deadlines, and the

like. 2) Self-induced stress is generally recognized as the self-imposition of a condition or stimulus leading to the type of arousal defined above. Examples include goals, deadlines, or performance requirements of any type with which the individual forces himself to function above a "comfortable" or "easy" level.

These distinctions often merge, however. In a prime example, Swedish military pilots were given a relatively complex set of procedures to execute in an unsophisticated flight simulator, and promised a reward for high performance. [Ref. 7] The situation was imposed on them, and could be construed as situational stress. The desire to perform at a high level, however, was self-induced. An interesting outcome of the experiment was that the physiological indicators of the pilots in question reacted remarkably like those of pilots in a similar non-simulated situation. A further interpretation of these studies was that subjects had come to accept the reality of the simulated situations, and their physiological indicators reflected as much. Such studies, in general, tend to support the validity of simulated situations in conducting stress research.

B. STRESS AND THE VOICE

Despite the abundance of general stress literature, relatively little has been done on how this phenomenon might

manifest itself in the human voice. Hicks [Ref. 8] gives a thoughtful overview of the recent literature, in addition to his own experimental results with various forms of stress and voice. He concedes the generally accepted notion that stress has an influence on individuals' fundamental speaking frequency (Fo); he concludes (as do others) that this parameter is as likely to go up as down, with concomitant frequency-range variations. The magnitude and direction of this shift in Fo is a function of the individual speaker. Additionally, the amount of shift required to constitute an indication of stress is still a subjective assessment. The one conclusive result, reported by Hicks and others [Ref. 8: p.99, p.124] is that the ratio of speaking time to total time will increase in a stressed speaker. This is consistent with everyday experience: "stressed" speakers often appear to talk in longer bursts, with shorter pauses separating them. Parameters other than the speech-to-pause ratio give inconclusive or contested results.

A large amount of commercial interest has focused on the influence of stress on vocal parameters. One can imagine the potential interest in being able to evaluate the stress level of a speaker. If done by voice, this analysis could be performed at great distances (via telephone, radio, etc.) and without the speaker's knowledge. Several firms have marketed devices purported to analyze the stress level of speakers (as a sort of voice polygraph), but their claims

appear subject to debate and are hotly contested in academic circles. These devices base their analysis on the existence of microtremor in the muscles controlling vocal activity, a phenomenon which remains to be shown conclusively. [Ref. 9: p.175]

Moreover, the readout provided by voice stress analyzers requires rather subjective interpretation. As noted, vocal variations so far demonstrated have shown a subject-specific nature: where one subject's Fo rises, another's may fall, etc. Thus, analysis of a speech sample for stress requires some apriori knowledge of the individual speaker's tendencies. Analytic efforts so far have been limited by the necessity for rather subjective analysis. Work is currently underway both to objectify such analysis and improve the analytical algorithms involved.[1]

In short, there currently no universal and few conclusive indicators of stress in a speaker's voice. It is widely accepted that psychological stress has an influence on vocal production, yet hotly contested in academic and industrial circles what the best indication of such stress is. In a military/applications context, it may well be that the most fruitful area of inquiry would be the implications of stress upon the voice, from the viewpoint of productivity and reliability with automatic speech recognizers.

[1]Conversations with Dr. D. Biers, University of Dayton, October 1982.

III. EXPERIMENTAL COMPONENTS

A. GENERAL

It is generally recognized to be important to simulate in an experiment the conditions under which the phenomena of interest might actually be expected to occur. This is no less important in experiments dealing with voice input, where there is a clearly definable task and one suspects an operative but unquantifiable factor such as stress. Moreover, with voice input, it is essential to use subjects representative of those who might actually be using ASR in the task-environment of interest.

B. SUBJECTS

Twenty-four volunteer subjects were obtained, mainly from the C3 curriculum at the Naval Postgraduate School. The makeup of the sample is shown below:

Table 1

SUBJECT DEMOGRAPHICS

	EXPERIENCED	INEXPERIENCED
MALE MILITARY	9	8
FEMALE MILITARY	3	1
MALE CIVILIAN	0	1
FEMALE CIVILIAN	0	2
TOTAL:	12	12

Twenty-one of these subjects were military officers, in the grade of O3 and O4, representing the four services; three were from civilian agencies. Six of the twenty-four were female. These subjects were at a level in their careers where, especially in light of their C3 background, they might reasonably expect to be at the mid- to upper levels of a Command-Post staff. In short, they were prime candidates for using voice input in a Command-Post scenario.

Of the twenty-four, half had some prior experience with ASR, which was defined for the purposes of the experiment as participation in three or more experiments at the Postgraduate School; the others had none. The sexes were evenly divided between the two groups. Batchellor [Ref. 10] found no statistically significant difference in recognition rates between sexes with the particular recognizer being used. The balanced sample (blocked on sex) was used nonetheless to control for any unforeseen effects due to gender.

C. VOCABULARY

For this experiment, subjects trained a fifty word vocabulary (Appendix A) which included some numerals and various other utterances. The vocabulary was chosen with a specific purpose in mind. It has been shown [Ref.10] that

correct recognition of a word or phrase is directly related to the length of the utterance. In this light, some words were deliberately chosen over others to avoid confusion or possible misrecognition. (For example "send message" in lieu of the shorter "send".) Other words were omitted for the same reasons. Some of the vocabulary items were used infrequently, but were necessary for the task. In sum, the vocabulary chosen was the minimum felt necessary to enable composing short messages with a modicum of realism and some variety.

I. TASK

Each subject was given a scenario consisting of 200 utterances (Appendix B), with each word or phrase numbered. These utterances, when correctly interpreted by the recognizer, brought up on the screen an output resembling messages which might be sent or serviced by a shipboard Command Post. (The ouput from these message sequences is shown in the right-hand column of Appendix B, which was not furnished to experiment participants; for examples of the types of messages see Figure 1.) Certain utterances produced a multiple-word output, simulating message-handling utilities such as are resident on the ARPANET, with which all subjects were familiar. Verisimilitude was an important consideration in the design of the task:

SAMPLE MESSAGE OUTPUT

```
SEND MSG TO:/ CV BATTLE GROUP / COMMANDER
INFO COPY TO: / COMTHIRDFLEET / CINCPACFLEET
FROM: / USS RATHBURNE
POSITION / REPORT / NUMBER / 1
READINESS LVL / 3
POSSIBLE / ELECTRONIC / SURFACE / CONTACT
BEARING: / 1/8/Ø / DISTANCE / 2/7
CONFIDENCE LVL/ 3
REQUEST / INSTRUCTIONS

QUIT./
```

```
FORWARD MSG #: / 7
INFO COPY TO: / CV BATTLE GROUP / COMMANDER
FROM: / USS RATHBURNE

QUIT./
```

(NB: The symbol / indicates the separation between utterances. Line separation is accomplished by speaking "RETURN" which yields a carriage return.)

Figure 1

SAMPLE MESSAGE OUTPUT

> The point of departure of the synthetic-work approach is a behavioral analysis of the performance requirements placed on the operator by some specific system, or by a class of such systems in general. Tasks are then selected against a criterion of content validity (i.e., tasks are selected because they measure functions judged by experts in the field to be important to the operational situation of interest). A general criterion of face validity is also imposed (i.e., the tasks are configured to be acceptable to target populations, e.g., pilots ... [Ref.11: pp. 22-23]

Subjects were to speak the utterances in order, each one until the proper output appeared on their screen. Each subject was given the same scenario three times in the course of the experiment, with at least one day (but no more than two) between reprises. On subsequent trials, subjects were told that the scenario they were given was of exactly the same length and of similar content to the one(s) they had seen previously. The same scenario was used repeatedly to control for variability of the task, and it was felt that the scenario was of sufficient length and variation to negate any significant learning effect.

E. TIME STRESS FACTOR

During the first phase of the experiment (Phase 1), subjects were allowed to complete the scenario at their own pace, and were exhorted not to rush or needlessly linger. (cf. Appendix C). The experimenter attempted to put them at ease. Unknown to the subjects, they were being timed, and their times to complete the scenario recorded. This time was

to serve as an individual baseline against which subsequent phases would be timed. In the next phases (Phases 2 and 3) subjects were allotted two-thirds and then one-third of their baseline times, respectively, to attempt to complete the scenario. It was anticipated that not all subjects would complete the scenario under the shortened time allowance: scoring was adjusted to reflect the amount of the task actually completed. (cf. IV.B infra)

The time-compression was intended to induce stress in the subjects. They were informed at the outset of how much time they would have to complete a task with which they were already familiar. A large Gra-Lab darkroom timer was plainly visible in front of the subjects to show time remaining. Additionally, the experimenter reminded the subjects insistently at thirty-second intervals of their time remaining. Subjects could guage their own progress in the 200 item scenario. Apriori, it was felt that the two-thirds time would constrain some but not all subjects; it was reasonably certain that allowing only one-third of the baseline time would do so. This opinion seemed to be borne out in pilot trials (with subjects not used in the later experiment).

F. EQUIPMENT AND SETUP

The voice recognition system employed was the T600 recognizer manufactured by Threshold Technology, Inc. (See

Figure 2 for a block diagram of the system.) This is a discrete utterance analyzer, which analyzes samples of speech up to a maximum of two seconds in length, separated by a distinct pause of at least one-tenth of a second. Utterances of more than one word are allowed, as long as the words are run together and pronounced "naturally", without distinct breaks between them.

The T600 has a maximum capacity of 256 utterances. In the training mode, the subject repeats a word or phrase ten times. (Under certain circumstances, fewer repetitions may be used, but this requires modification of the software and yields less consistent results [Ref. 10]). After the tenth repetition, the ten are superimposed and normalized in length to form a template for that utterance. During this experiment, a training standardization measure was adopted. When the subjects had finished training, they repeated the vocabulary list in toto at least twice; any items not correctly recognized on at least two of three passes were retrained.

When the user speaks, his speech sample is compared against all these templates for the best match. If within preset tolerances, the utterance is recognized (correctly or not) and the output string sent to the display screen. If no match is found, the machine produces an audible "beep". Each individual may store his voice template on a data-tape cartridge for future use, and thus bypass the lengthy

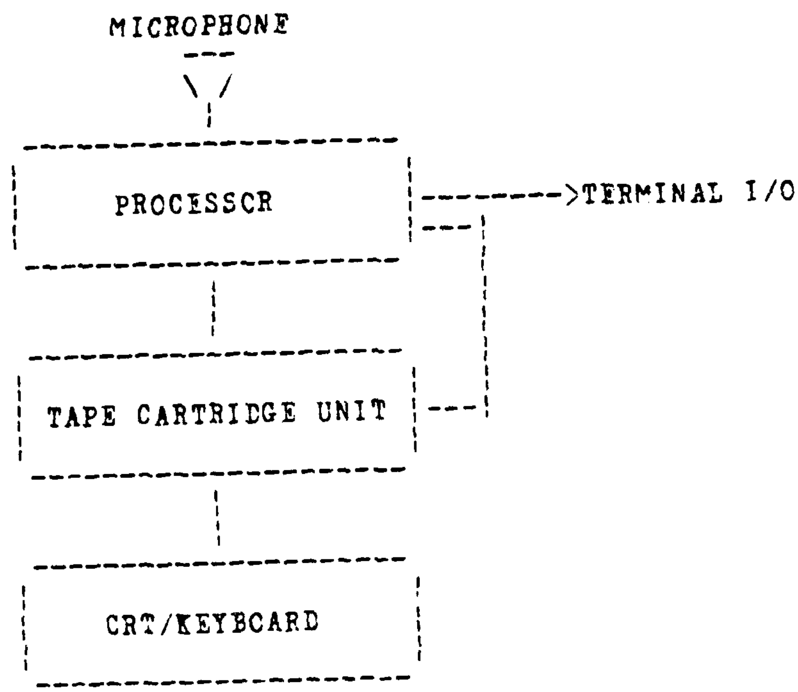


Figure 2

T600/CRT SET-UP BLOCK DIAGRAM

training process in future sessions. Each utterance is programmed by the user (in this case, the experimenter) with a prompt for training. When that utterance is cued at the keyboard, the prompt-string appears on the display; it disappears after the tenth repetition. After training, the recognizer shifts to the recogniton mode and awaits verbal input. When utterances are recognized, an associated preprogrammed output string is sent to the display screen. The output string corresponding to an utterance is user-selected, allowing up to a maximum of sixteen alphanumeric characters.

When used in conjunction with a computer, the T600 may be used in either direct or buffered modes. "Direct" sends the output string to the computer without delay. The "Buffered" mode allows the user to verify the output before allowing it to be sent. The direct mode was used here. In this experiment, the T600 was used in a stand-alone capacity: output strings were sent to the display screen and nowhere else. (For further discussion see [Ref. 12])

The microphone used was a Shure SM10 noise cancelling unit. It was worn by all subjects on the left side of the head, directly below and slightly to the left of center of the lips. To further attenuate echo and outside noise, the experiment was conducted inside an Industrial Acoustics Company soundproof chamber. (See Figure 3 for the actual equipment setup).

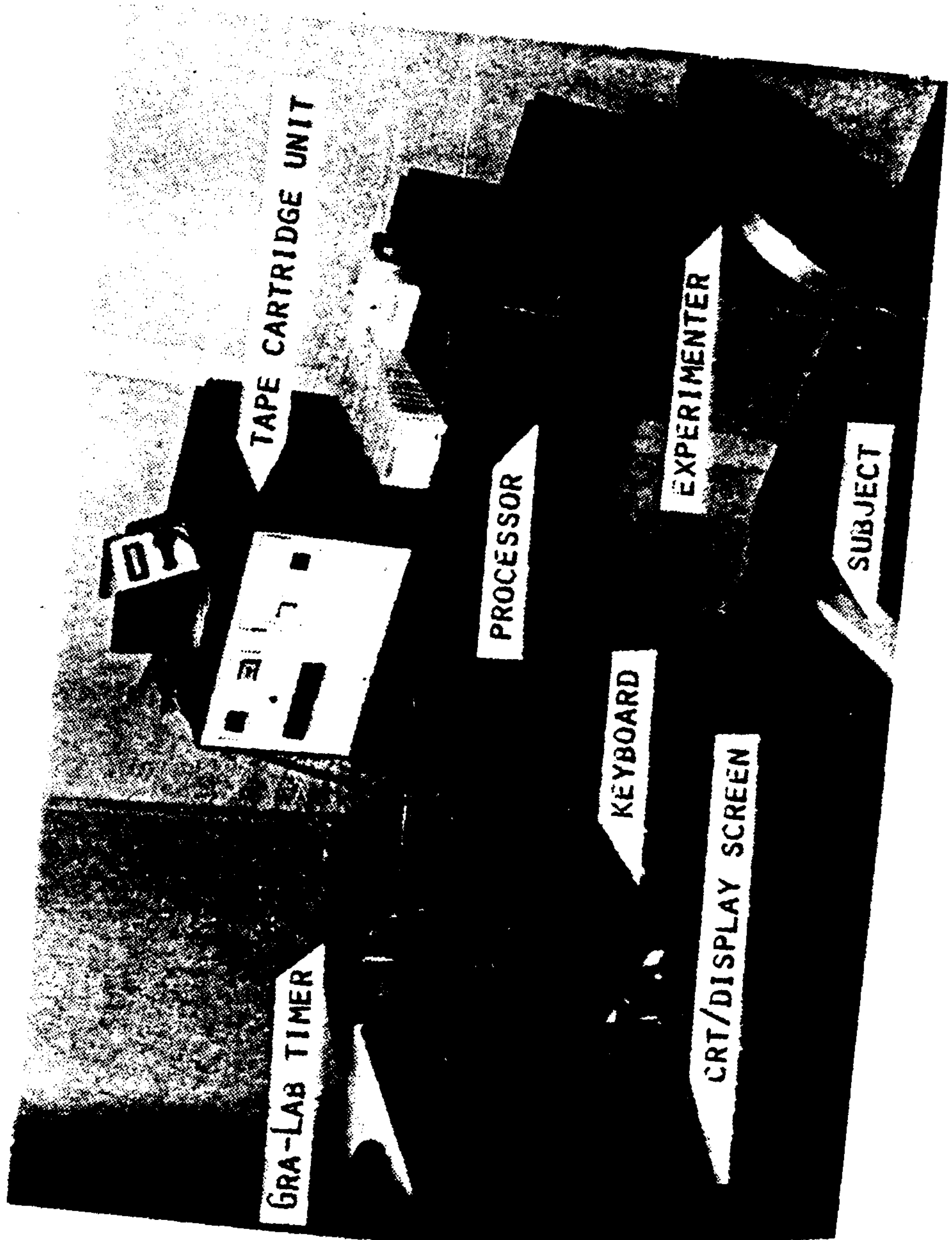


Figure 3

EQUIPMENT SETUP

IV. EXPERIMENTAL DESIGN

A. OBJECTIVE

The primary objective was to see if time stress affected users' recognition rates with this voice recognition setup. After the initial training session, (and after a delay of three to four days) subjects began the portion of the experiment which involved the scenario. They were fitted with the microphone and handed the scenario. For each respective phase, they were read the instructions in Appendices C,D, and E. The intention was to induce stress in subjects in the last two phases of the experiment by progressively shortening the time allowed to perform a task of fixed length. Compression of time allowed was to be correlated with percent correct recognition rate.

It was hoped that several factors were to be at work on the participants. There was a definite competition for the subjects' mental resources. Subjects had to read a word or phrase from their printed scenarios, speak it, and verify that it appeared correctly on their screen. If misrecognized or "beeped", they were to try again until correctly recognized. The experimenter instructed subjects to proceed to the next scenario item if they had five unsuccessful attempts, although they had no foreknowledge of this. This was yet another factor in the stress equation: failure at a

task is in itself stressful for most individuals.[Ref. 13: p.151]

At the outset of each of the last two (time-constrained) phases of the experiment, subjects were informed of how much time they would be allotted to perform the scenario. After the second and before the last phase, many indicated to the experimenter that they "knew" they would have less time than in the previous session. This suspicion was desired and was calculated to create a measure of negative anticipation in the participants. During the last two phases of the experiment, subjects were reminded insistently by the experimenter of time remaining, in addition to the presence of the highly visible timer. It was estimated that this combination of factors would induce stress:

> The apprehension of psychological stressors is a much more complex business. It depends on: (1) prior experience with a class of stressors; (2) remembering the experience; (3) recognizing that the present stimulus gestalt is an instance of the earlier class of stressors; and (4) believing that the likelihood of the threatened occurrence is above some subjective threshhold. [Ref. 13: p. 142]

Moreover,

> While the experimental manipulation allows a fairly high degree of control over the stimulus conditions, the reaction patterns of subjects are often difficult to determine. ... Thus, the stress level induced may be minimal for some subjects, whereas it may be moderate or rather strong for others. Since it is often difficult to determine a definite external criterion for stress (even psychophysiological variables cannot always be expected to indicate the stress level of an individual consistently and validly,) experimenters using laboratory stress induction can only hope that they have

been able to induce a reasonable amount of stress in their subjects. [Ref. 9: p. 176]

The fundamental hypothesis being tested here may be summed up as:

Ho: Time stress has no effect on the recognition rates of users of voice recognition systems.

H1: Time stress affects the recognition rates of users of voice recognition systems.

Additionally, it was hoped to see whether:

Ho: Experienced and inexperienced users of voice recognition systems are affected equally by time stress.

H1: Experienced and inexperienced users of voice recognition systems are not affected equally by time stress.

B. SCORING

Subjects were monitored during their performance of the scenario: the experimenter sat to the right of and behind subjects (cf. Figure 3) from which position both the subjects' actions and the display screen were clearly visible. Misrecognitions (improper word substitutions by the recognizer) and non-recognitions were recorded. A confusion matrix of substitutions and "beeps" is included as Appendix

8. Subjects occasionally lost their place in the scenario, repeating some vocabulary items they had already spoken. These repetitions were counted as additional utterances, and the correct or incorrect result recorded; thus, no penalty was given for repetition. The caveat in the instructions about penalties for non-completion of the scenario was bogus. It was added to provide an additional incentive to finish, and in the hope of reinforcing the subjects' apprehension. (Subjects completed an average of fifty-three percent of the scenario in Phase 3; none finished.) Recognition rates were based on the items of the scenario actually completed. Percentage correct recognition was figured as the quotient of the number of correct recognitions divided by the total number of utterances spoken by the subjects (including mis- and non-recognitions).

At the end of each phase of the experiment, subjects were asked to rate their feelings during that phase. (See questionnaire, Appendix G). Five surrogate descriptors were used to circumscribe the idea of "stressed"; subjects were asked if they felt:

1. nervous

2. confused

3. pressured

4. irritated

5. anxious

on a scale of one to five, five being the highest level of each. It was thought that these individual response levels might somehow be related to recognition rates.

C. CONCEPTUAL DESIGN OF THE EXPERIMENT

The design employed was a two-factor mixed design with repeated measures on one factor. [Ref. 14: pp. 54-61] Alternatively (and for purposes of clarity) the design might be thought of as involving two conditions (experienced/inexperienced) and three treatments. (See Figure 4 for a graphic interpretation).

As stated previously, the treatments consisted of three variations of time in the testing, constituting the three phases (excluding the training phase).

1. Phase 1	baseline time: no constraint
2. Phase 2	2/3 baseline time
3. Phase 3	1/3 baseline time

These phases were conducted in the order shown. While it

TIME FACTOR --->

EXPERIENCE LEVEL	PHASE 1 (BASELINE)	PHASE 2 (2/3 BASELINE)	PHASE 3 (1/3 BASELINE)
EXPERIENCED	SUBJECT 1 . . . . . SUBJECT 12		
INEXPERIENCED	SUBJECT 13 . . . . . SUBJECT 24		

Figure 4

CONCEPTUAL DESIGN OF THE EXPERIMENT

might be thought preferable in the classical design of experiments to randomize the administration of the time constrained portions, this was not done. Randomization might have controlled for the possible effects of "learning" between Phases 2 and 3; this benefit was consciously sacrificed in order to introduce the apprehension factor discussed above. It was suspected apriori that the drastic time compression from the second to the third phase of the experiment would help to compensate for the possible confounding introduced by "learning". Moreover, the scenario was felt to be long and varied enough to counteract learning in some measure, in addition to the experimenter's deception that successive scenarios were not exactly the same. The fact that all phases were separated by no less than a day (and no more than two) was also felt to counteract learning.

V. RESULTS AND ANALYSIS

A. ANALYSIS

The results of the experiment as described are shown graphically in Figure 5. Before conducting any analysis, Hartley's test for equality of variance was performed on the raw data. The results indicated that equality of variance was an invalid assumption, and thus the raw data required transformation: after arcsin transformation, the equality of variances was accepted. (See Figure 6)

The analysis of variance table is presented in Figure 7. As is readily apparent from the graph of the raw data, recognition rates for the experienced group decrease monotonically with time compression. This is also true for Phases 2 and 3 of the inexperienced group. The F statistic shows a significant difference between trials. Therefore, the null hypothesis of time stress not affecting recognition rate is rejected.

An interesting phenomenon is observed with the inexperienced group, however. From Phase 1 to Phase 2 their success rate increases noticeably, despite the decreased time allotted for the task. This is hypothesized to be the result of a combination of factors, the most obvious of which is learning. From their first, open-ended trial to the second, the inexperienced group became more adept at the task. Moreover, given their relative naivete with the

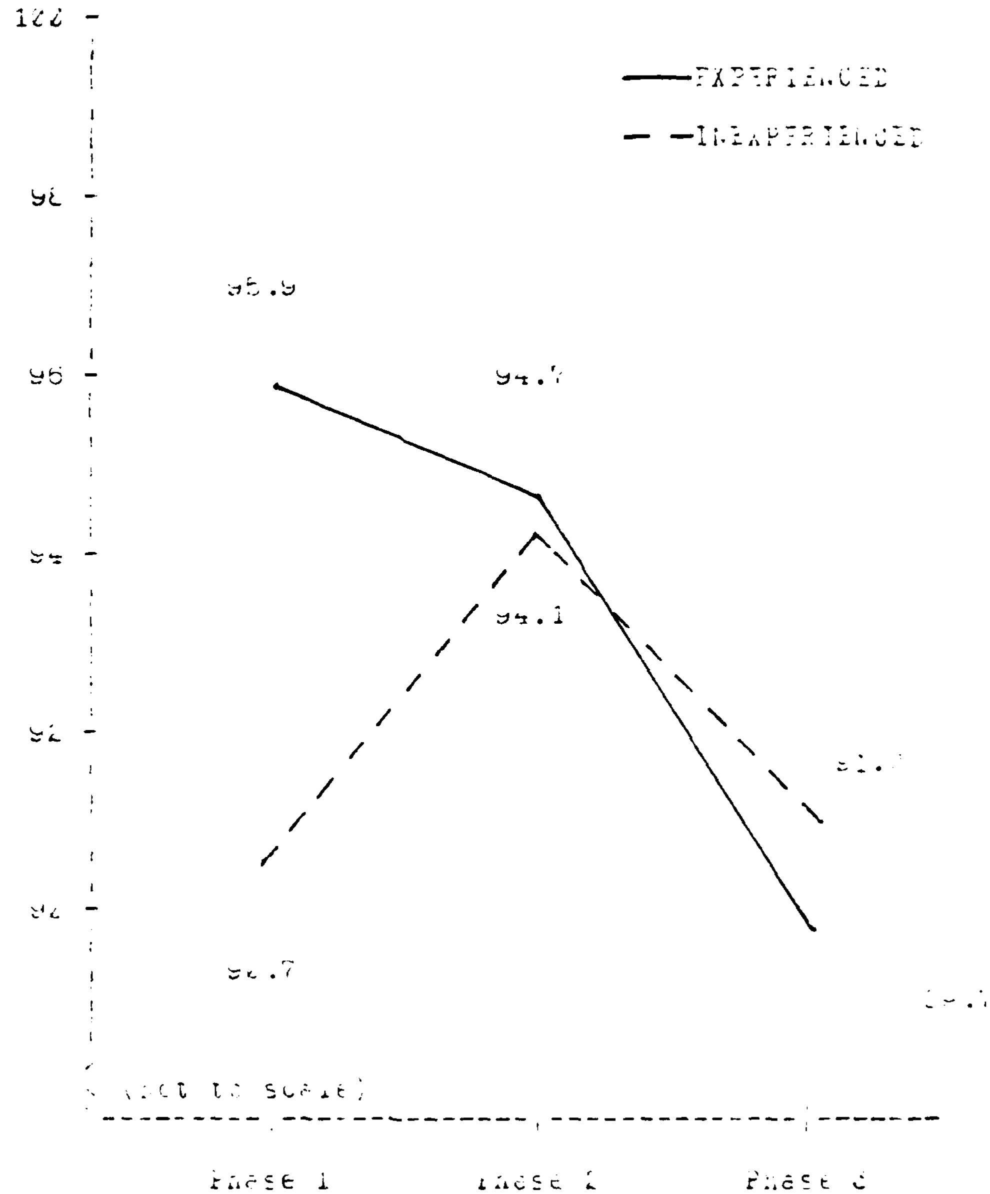


Figure 5

PERCENTAGE CORRECT RECOGNITION VERSUS EXPERIMENT PHASE

PRELIMINARY RESULTS

SUMMARY OF EXPERIMENTAL RESULTS

RAW DATA
(For percentages, multiply by 100)

		$\overline{X}$	$S^2_{\overline{X}}$
	Phase 1	.959	.0033
EXPERIENCED	Phase 2	.947	.0012
	Phase 3	.897	.0041
	Phase 1	.907	.0081
INEXPERIENCED	Phase 2	.941	.0023
	Phase 3	.912	.0037

For df1=t=6 and df2=(n-1)=11,
Hartley's test, Fmax = 6.32 (alpha=.05).
For the untransformed data:

$$S^2max/S^2min = .0081/.0012 = 6.75$$

Therefore, reject equality of variances; transform data.

After the transformation ($y = 2(\arcsin\sqrt{x}\ )$):

$$S^2max/S^2min = .0428/.0121 = 3.537$$

Accept equality of variances.

Figure 6

DATA SUMMARY AND HARTLEY'S TEST

SOURCE	SS	df	ms	F	p
Total	4.292	71			
Between subjects	2.074	23			
Experience	.0642	1	.0642	.722	
Error (b)	2.041	22	.0914		
Within subjects	2.218	48			
Time Factor	.3242	2	.1621	4.27	$<.05$
TimexExperience	.2216	2	.1108	2.92	$.05<p<.10$
Error (w)	1.672	44	.038		

Figure 7

ANALYSIS OF VARIANCE TABLE

equipment setup and the task, inexperienced subjects took longer to complete the initial phase of the experiment and thus had a longer baseline (an average of 7.89 minutes, versus 6.48 for the experienced). In the linear compression of Phases 2 and 3, they simply had more time than the experienced group.

It was not known apriori if allowing two-thirds of baseline time in the second phase would impose a felt time-constraint on all subjects. For this reason, one-third was chosen for the final phase: although seemingly impossible, it was felt that such a drastic time reduction would induce stress in the majority of the subjects. Such appears to be the case. Experienced subjects show a decline in each successive phase of the experiment. Inexperienced subjects, although theoretically still on the upward side of their "learning curve" show a pronounced decline from Phase 2 to Phase 3.

For the inexperienced group, both the rapidity of the learning effect and the quickness with which it was negated have additional significance. A Newman-Keuls range test (alpha = .05) was performed for differences in the means amongst phases of the experiment. [Ref. 16: pp. 35-37] Phases 1 and 2 were seen to be significantly different from Phase 3, but were not differentiable between themselves: this is obviously due to the inexperienced group starting at a lower rate, then coming up to meet the experienced group's

mean score. A two-sample "t"-test shows a difference between the experienced and inexperienced groups only in the first phase. (At an alpha critical of .069). In subsequent phases of the experiment, the two groups are statistically indistinguishable. This is further confirmed by the F statistic for experience in the analysis of variance. The inference to be drawn from this appears to be that the effects of learning on this type of equipment are so rapid as to make the two groups quickly very similar.

B. CORRELATION

First, an attempt was made to correlate subjects' questionnaire responses with the phase of the experiment to which they were responding. It was desired to see if response levels were somehow dependent on the time stress levels imposed on the participants. If so, a positive correlation would show a rising response level to one or more of the five questions as time compression increased with experiment phase. The data to be correlated were nominal (phase) versus ordinal (1 to 5 on questionnaires) so the nonparametric Spearman's Rho was used. [Ref. 17: pp. 252-256] Table 2 shows the results:

Table 2

Correlation of Questions with Phase

QUESTION	RHO
1. nervous	.234 *
2. confused	.017
3. pressured	.533 *
4. irritated	.020
5. anxious	.230 *

(* indicates significant at alpha less than .05)

Thus, three of the five questions show a strong correlation with the time compression with phase. Moreover, these three particular descriptors (nervous, pressured, and anxious) were thought to be the closest surrogates of the five to otherwise describe the situational or time stress desired. (The other two show no correlation.) It appears that the experimental conditions were successful in inducing stress in some measure. Mean response rates for subjects' answers are shown in Figures 8-12.

Next, an attempt was made to correlate subjects' questionnaire responses (on the five descriptors already mentioned) to their recognition rates with the system. It was suspected beforehand that response levels to one or more of these questions might increase monotonically from one

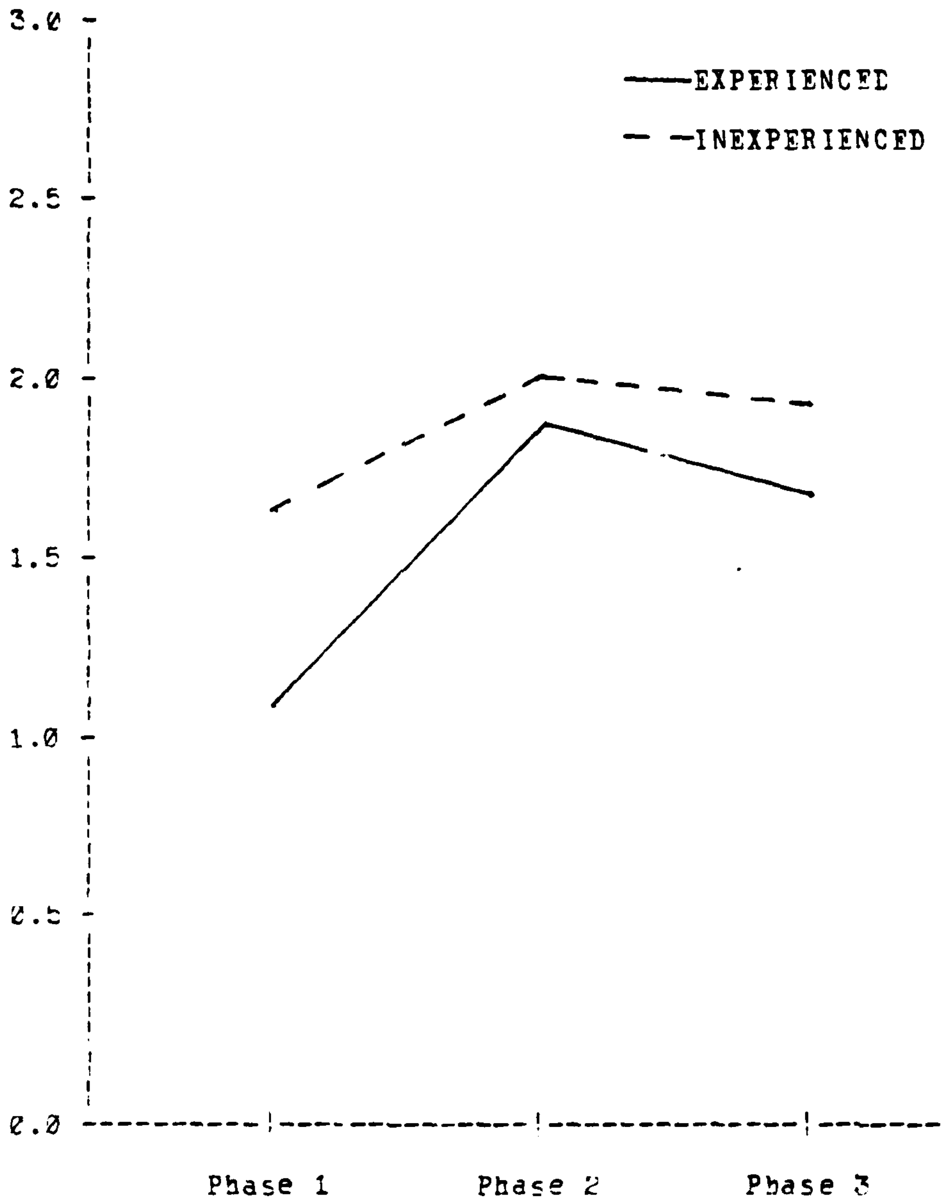


Figure 8

MEAN RESPONSE LEVELS TO QUESTION 1, "NERVOUS"

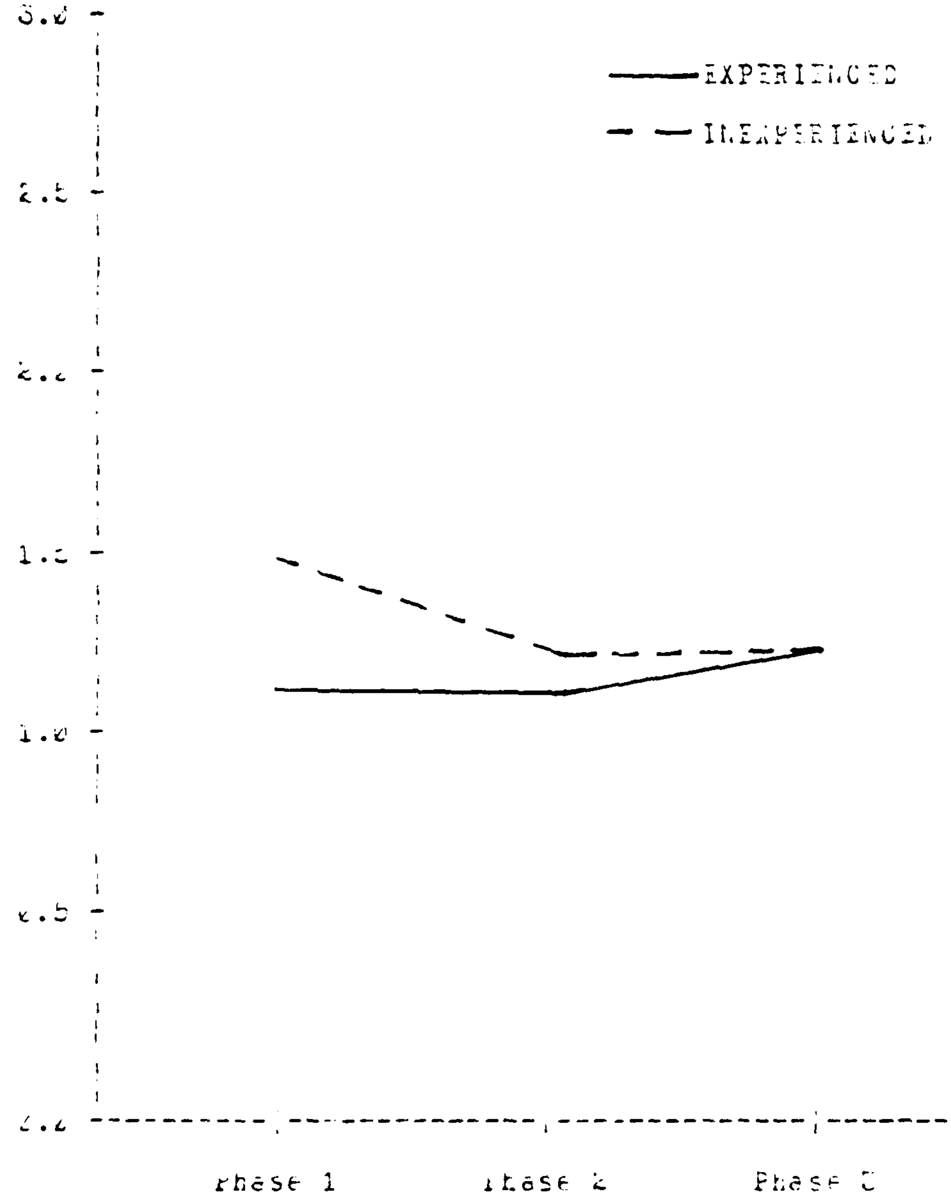


Figure 9

MEAN RESPONSE LEVELS TO QUESTION 2, "CONFUSED"

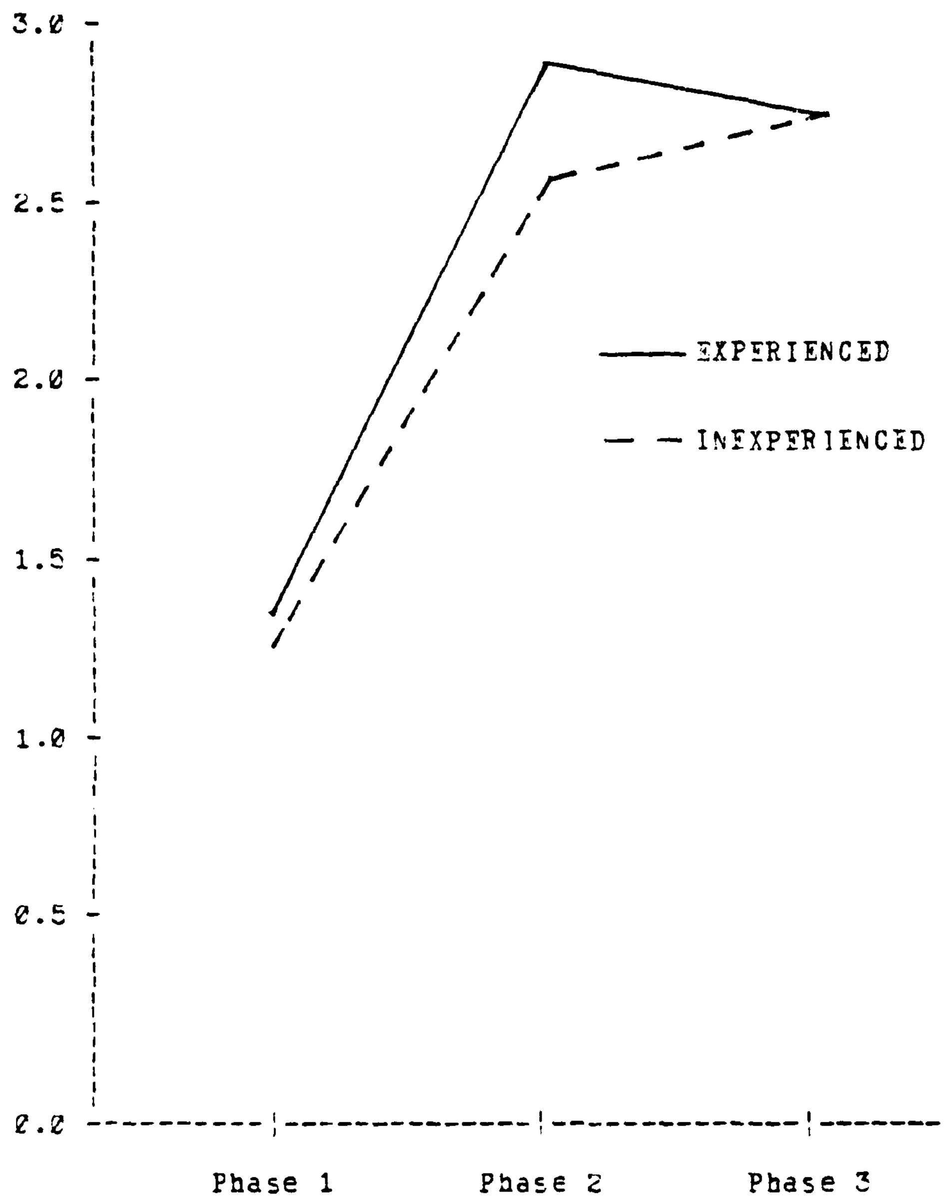


Figure 10

MEAN RESPONSE LEVELS TO QUESTION 3, "PRESSURED"

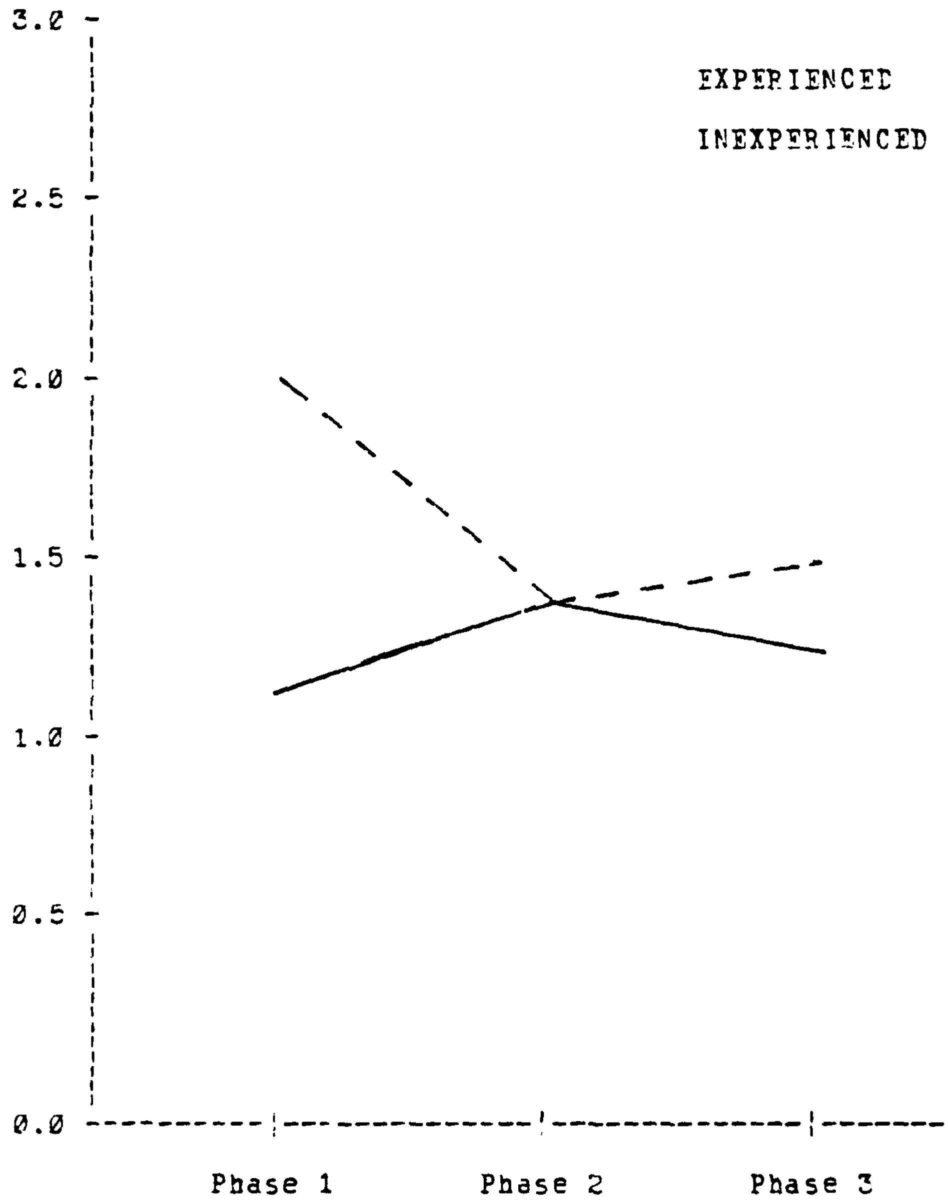


Figure 11

MEAN RESPONSE LEVELS TO QUESTION 4, "IRRITATED"

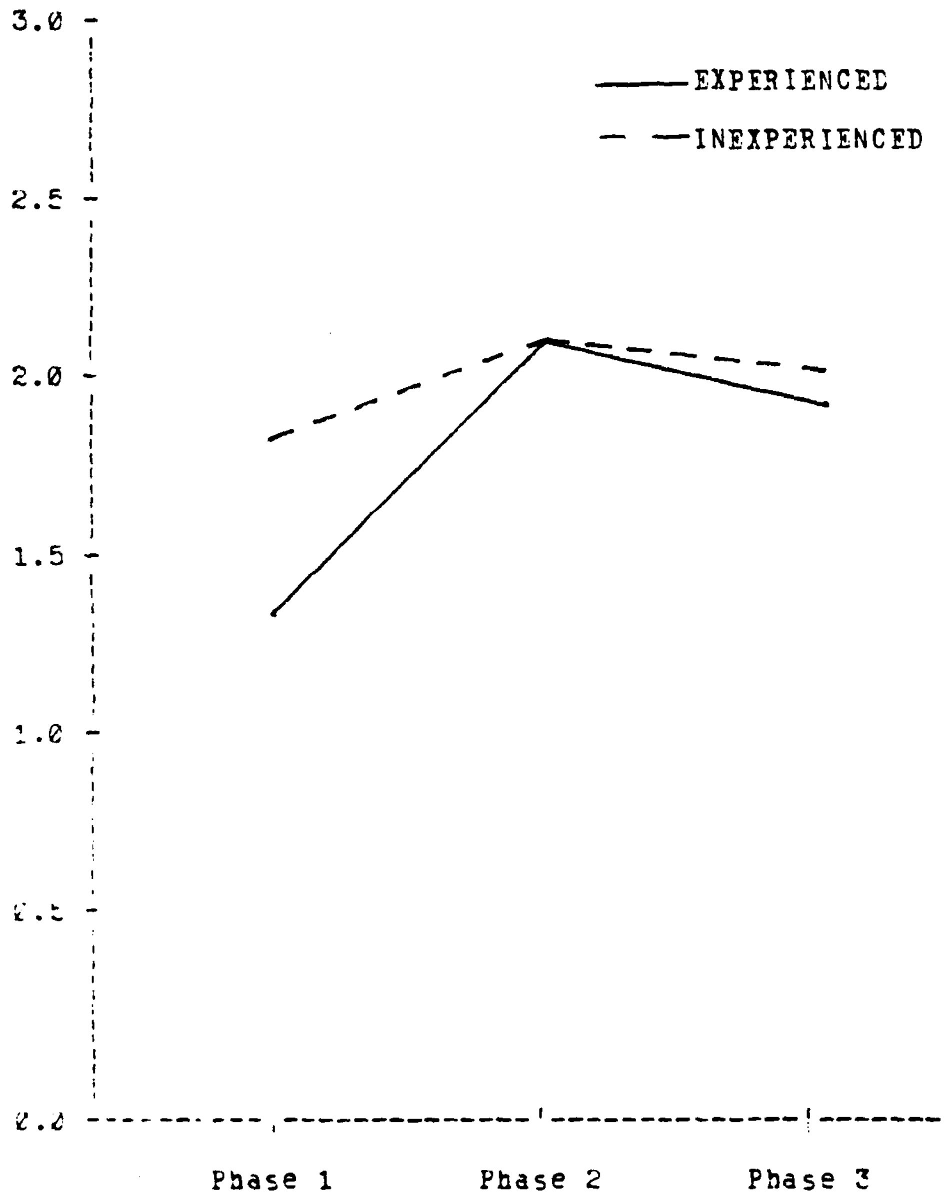


Figure 12

MEAN RESPONSE LEVELS TO QUESTION 5, "ANXIOUS"

phase to the next during the conduct of the experiment. It is readily apparent that this was not the case. While the five surrogate descriptors were deemed capable, in the aggregate, to describe stress, they apparently were not completely successful as a subjective measure.

The risk of self-inhibition is always run when asking subjects to rate their own perceptions, emotions, etc. [Ref. 18: pp. 501-503] Self-image or self-esteem may prohibit a respondent from admitting of a higher level of what he may feel to be an "undesirable" feeling or trait. This may have been the case here. In the course of the experiment, the experimenter's subjective assessment of the subjects' emotional states did not agree well with their questionnaire responses. On several occasions, subjects left the experiment giving every appearance of being agitated and exercised, having answered with a "1" (i.e. not at all) to all five questions.

Moreover, although instructed to respond only in regard to each individual phase of the experiment, it appears likely that subjects' responses may have been in the context of their previous experiences in the experiment. Thus one sees the rise in most responses from the first (untimed) to the second phase of the experiment. During the final phase, subjects were expecting to be placed under a tighter time constraint (the desired apprehension effect). This may in

part explain the absence of a rise in response levels for this last phase.

The data gathered from the questionnaires was in the ordinal (1 to 5) scale. Since an attempt was to be made to correlate these responses with recognition rates (interval data) recourse was again made to the non-parametric Spearman's Rho procedure for correlating the ranks of both sets of data. Recognition rates were tested against responses to all five questions for each phase of the experiment. Results generally showed no correlation.

One question alone provided a rather consistent negative correlation with recognition rates. For the inexperienced subjects, question 2 ("confused") showed a strong negative correlation (rho< -0.4) for all three phases of the experiment (credible at an alpha less than or equal to 0.1). This would imply that as confusion level rose, recognition rate went down. It is interesting to note that there was no correlation found between this variable and recognition for the experienced group. The only other question which showed some (negative) correlation was the fourth ("irritated"), but this relation was at most tenuous (0.25>alpha>0.1). Also of note is the fact that, after the initial phase of the experiment, response levels for the two groups parallel one another, in some instances very closely. This would appear further to corroborate the hypothesis that differences between the two groups diminish rapidly.

VI. REMARKS

If psychological stress has an effect on the voice, as is generally acknowledged, it is reasonable to expect it to have an effect on the success rates of users of voice input equipment. As discussed previously, ASR technology is of potentially great interest to the military planner in the Command, Control, and Communications context. This preliminary inquiry has attempted to show a correlation between stress and recognition success with such equipment. Several areas of consideration suggest themselves.

As has been seen, recogniton rates declined when the operators of this ASR equipment were placed under time stress. An alternative approach to evaluating the desirability of this technology might be to assess voice versus other means of input under similar conditions. In this study, subjects attempting to complete a nearly impossible task (Phase 3) nonetheless managed to maintain a correct recognition rate of approximately ninety percent. With manual input (e.g., typing), each characer to be input represents an opportunity for error: how this might be affected by psychological stress is open to speculation. When one considers that a sixteen character output string can be produced with one voice command, it may well be that

voice input becomes even more desirable in a stressful environment.

It is interesting to note how quickly the inexperienced subjects "caught up" with and paralleled the performance of their experienced counterparts. If this is indeed a generalizable phenomenon, it would imply that after some few training sessions with a recognizer, the distinction vanishes. If so, faced with a potentially adverse situation a commander or watch officer need not worry about this experience-level in his Command Post staffing. Voice input has also been proposed for automated reservation and directory-assistance schemes. Given the success rates for inexperienced users even under conditions calculated to induce stress, these applications appear no less feasible.

Lastly, it is generally acknowledged that a more objective method for assessing psychological stress is desperately needed. Any further work in this area should find an alternative and hopefully more workable method of assessing stress in subjects. It is submitted that some relatively non-intrusive physiological measures of stress be considered. Several of these observable and quantifiable indicators are well documented in existing stress literature.

APPENDIX A

****
VOCABULARY LIST
****

#	WORD
0.	ZERO
1.	ONE
2.	TWO
3.	THREE
4.	FOUR
5.	FIVE
6.	SIX
7.	SEVEN
8.	EIGHT
9.	NEGATIVE
10.	TERMINATE
11.	SEND MESSAGE
12.	ANSWER MESSAGE
13.	FORWARD MESSAGE
14.	CINCPACFLEET
15.	THIRDFLEET
16.	BATTLE GROUP
17.	INFORMATION
18.	RETURN
19.	RADAR
20.	VISUAL
21.	ELECTRONIC
22.	CONTACT
23.	SIGHTING
24.	SOVIET
25.	UNKNOWN
26.	POSSIBLE
27.	CERTAIN
28.	CONFIDENCE
29.	DAMAGE
30.	ASSESSMENT
31.	REQUEST
32.	INSTRUCTIONS
33.	READINESS
34.	INTENT
35.	ATTACK
36.	REPORT
37.	RATHBURNE
38.	FROM
39.	COMMANDER
40.	SURFACE
41.	TIME
42.	ZULU
43.	PERIOD
44.	POSITION
45.	NUMBER
46.	BEARING
47.	DISTANCE
48.	UNDER
49.	ADVISE

APPENDIX B

****
SCENARIO LISTING   CRT OUTPUT
****

1. SEND MESSAGE
2. BATTLE GROUP
3. COMMANDER
4. RETURN
5. INFORMATION
6. THIRDFLEET
7. CINCPACFLEET
8. RETURN
9. FROM
10. RATHBURNE
11. RETURN
12. POSITON
13. REPORT
14. NUMBER
15. ONE
16. RETURN
17. READINESS
18. THREE
19. RETURN
20. POSSIBLE
21. ELECTRONIC

```
SEND MSG TO: CV BATTLE GROUP COMMANDER
INFO COPY TO: COMTHIRDFLEET CINCPACFLEET
FROM: USS RATHBURNE
POSITION REPORT NUMBER 1
READINESS LVL 3
POSSIBLE ELECTRONIC SURFACE CONTACT
BEARING:180 DISTANCE 27
CONFIDENCE LVL 3
REQUEST INSTRUCTIONS

QUIT.
```

22. SURFACE

23. CONTACT

24. RETURN

25. BEARING

26. ONE

27. EIGHT

28. ZERO

29. DISTANCE

30. TWO

31. SEVEN

32. RETURN

33. CONFIDENCE

34. THREE

35. RETURN

36. REQUEST

37. INSTRUCTIONS

38. TERMINATE

39. FORWARD MESSAGE

40. SEVEN

41. RETURN

42. INFORMATION

43. BATTLE GROUP

44. COMMANDER

45. RETURN

```
FORWARD MSG #:7
INFO COPY TO: CV BATTLE GROUP COMMANDER
FROM: USS RATHBURNE

QUIT.
```

46. FROM

47. RATHBURNE

48. TERMINATE

49. ANSWER MESSAGE

50. TWO

51. FOUR

52. RETURN

53. CONTACT

54. UNKNOWN

55. POSSIBLE

56. SOVIET

57. PERIOD

58. INTENT

59. UNKNOWN

60. TERMINATE

ANSWER MSG #: 24
CONTACT UNKNOWN POSSIBLE SOVIET.
(no carriage return in scenario)
INTENT UNKNOWN

QUIT.

61. SEND MESSAGE

62. BATTLE GROUP

63. COMMANDER

64. RETURN

65. INFORMATION

66. THIRDFLEET

67. CINCPACFLEET

68. RETURN

69. CONTACT
70. SIGHTING
71. REPORT
72. NUMBER
73. SIX
74. RETURN
75. TIME
76. ONE
77. SEVEN
78. FOUR
79. EIGHT
80. ZULU
81. RETURN
82. POSITION
83. ONE
84. FIVE
85. SEVEN
86. TWO
87. ZERO
88. RETURN
89. COMMANDER
90. ASSESSMENT
91. POSSIBLE
92. SOVIET
93. SURFACE

```
SEND MSG TO: CV BATTLE GROUP COMMANDER
INFO COPY TO: COMTHIRDFLEET CINCPACFLEET
CONTACT SIGHTING REPORT NUMBER 6
TIME 1748 ZULU
POSITION 15720
COMMANDER ASSESSMENT POSSIBLE SOVIET SURFACE.
(no carriage return in scenario)
UNKNOWN

QUIT.
```

94. PERIOD

95. UNKNOWN

96. TERMINATE

97. ANSWER MESSAGE

98. THREE

99. EIGHT

100. RETURN

101. INFORMATION

102. CINCPACFLEET

103. RETURN

104. FROM

105. RATHBURNE

106. RETURN

107. UNKNOWN

108. SURFACE

109. CONTACT

110. PERIOD

111. POSSIBLE

112. SOVIET

113. RETURN

114. UNKNOWN

115. RADAR

116. CONTACT

117. BEARING

118. TWO

```
ANSWER MSG #: 38
INFO COPY TO: CINCPACFLEET
FROM: USS RATHBURNE
UNKNOWN SURFACE CONTACT. POSSIBLE SOVIET
UNKNOWN RADAR CONTACT BEARING: 235
DISTANCE 14
NEGATIVE VISUAL CONTACT.
CONFIDENCE LVL ASSESSMENT 6.
INTENT UNKNOWN. REQUEST ADVISE

QUIT.
```

119. THREE

120. FIVE

121. RETURN

122. DISTANCE

123. ONE

124. FOUR

125. RETURN

126. NEGATIVE

127. VISUAL

128. CONTACT

129. PERIOD

130. RETURN

131. CONFIDENCE

132. ASSESSMENT

133. SIX

134. PERIOD

135. RETURN

136. INTENT

137. UNKNOWN

138. PERIOD

139. REQUEST

140. ADVISE

141. TERMINATE

142. SEND MESSAGE

143. BATTLE GROUP

144. COMMANDER

145. RETURN

146. INFORMATION

147. THIRDFLEET

148. RETURN

149. FROM

150. RATHBURNE

151. RETURN

152. CONTACT

153. REPORT

154. NUMBER

155. SIX

156. RETURN

157. CONTACT

158. SIGHTING

159. RETURN

160. CONTACT

161. CERTAIN

162. SOVIET

163. SURFACE

164. RETURN

165. BEAPING

166. TWO

167. THREE

SEND MSG TO: CV BATTLE GROUP COMMANDER
INFO COPY TO: COMTHIRDFLEET
FROM: USS RATHBURNE
CONTACT REPORT NUMBER 6
CONTACT SIGHTING
CONTACT CERTAIN SOVIET SURFACE
BEARING: 237
DISTANCE 14
INTENT UNKNOWN

QUIT.

168. SEVEN

169. RETURN

170. DISTANCE

171. ONE

172. FOUR

173. RETURN

174. INTENT

175. UNKNOWN

176. TERMINATE

177. SEND MESSAGE

178. THIRDFLEET

179. RETURN

180. INFORMATION

181. CINCPACFLEET

182. RETURN

183. FROM

184. RATHBURNE

185. RETURN

186. COMMANDER

187. ASSESSMENT

188. BATTLE GROUP

189. UNDER

190. SOVIET

191. ATTACK

```
SEND MSG TO: COMTHIRDFLEET
INFO COPY TO: CINCPACFLEET
FROM: USS RATHBURNE
COMMANDER ASSESSMENT CV BATTLE GROUP
(no carriage return in scenario)
UNDER SOVIET ATTACK. DAMAGE LVL
(no carriage return in scenario)
UNKNOWN. REQUEST ADVISE USS RATHBURNE
(no carriage return in scenario)
INTENT

QUIT.
```

192. PERIOD

193. DAMAGE

194. UNKNOWN

195. PERIOD

196. REQUEST

197. ADVISE

198. RATHBURNE

199. INTENT

200. TERMINATE

****
END SCENARIO
****

APPENDIX C

INSTRUCTIONS TO EXPERIMENT PARTICIPANTS *PHASE 1*

The experiment you will be participating in simulates what might happen in a shipboard command post. You will be using the vocabulary you have already trained during the first session.

You will be given a list of words to speak to the machine: when recognized, they will come up on your screen. Certain words produce a multiple word output, as you have already seen. These are an attempt to simulate message-handling utilities such as are found on the ARPANET.

Speak clearly and distinctly, but don't overenunciate. No provision has been made to let you back up and correct mistakes, misrecognitions, etc., so you must continue to say the word until it is recognized. If you miss a word and don't catch it, I will prompt you for the word you missed; please simply try again on the word I prompt you with. Pauses between words may be short, but they must be distinct.

Set your own pace,don't rush but don't linger. You are not competing against anyone (including yourself) for time or recognition, so don't be nervous. No grades are being given out here. Once you start, however, please continue to the end.

I will be glad to answer any questions you may have about the conduct of this phase of the experiment. Any questions about the purpose and outcome I will be glad to discuss at great length once the actual experimentation phase is through.

Thanks again for your help.

APPENDIX D

INSTRUCTIONS TO EXPERIMENT PARTICIPANTS *PHASE 2*

This is Phase 2 of the experiment. In this phase you will participate in a Command Post scenario very similar in content and length to the first scenario you were given.

The most significant difference in this phase is the time factor. In Phase 2 you will be allotted only 2/3 of the time you originally took to complete the entire scenario. In fact, you may not be able to complete the entire scenario. It is to be emphasized that your primary objective is to complete as much of the scenario as you possibly can, as accurately as possible. In this command post situation, the messages you will be handling are of an urgent nature, and your primary motivation is to get as many of them completed as possible. You are being scored for both recognition rate and the percentage of the scenario you complete. Any parts you do not com- plete will count against your overall score, unfortunately....

During the experiment, I will give you frequent reminders of the time you have remaining at 30 second intervals. The time I announce will be in minutes and seconds remaining. The timer will also furnish you with a running indication of time remaining.

Again, attempt to speak clearly and distinctly. No provision has been made to back up and correct mistakes or misrecognitions, so you must repeat the word or phrase until it is correctly recognized. If you miss a word and don't catch it, I will prompt you for the word you missed: please try again on the word I prompt you with.

Let me re-emphasize here that the shortened time-allowance is meant to represent the urgency of the situation in the Command Post, and that your principal goal is to finish as much of the scenario as possible.

Thanks again for your help.

APPENDIX E

INSTRUCTIONS TO EXPERIMENT PARTICIPANTS *PHASE 3*

This is Phase 3 of the experiment. In this phase you will participate in a Command Post scenario very similar in content and length to the first scenario you were given.

The most significant difference in this phase is the time factor. In Phase 3 you will be allotted only 1/3 of the time you originally took to complete the entire scenario. In fact, you may not be able to complete the entire scenario. It is to be emphasized that your primary objective is to complete as much of the scenario as you possibly can, as accurately as possible. In this command post situation, the messages you will be handling are of an urgent nature, and your primary motivation is to get as many of them completed as possible. You are being scored for both recognition rate and the percentage of the scenario you complete. Any parts you do not com- plete will count against your overall score, unfortunately....

During the experiment, I will give you frequent reminders of the time you have remaining at 30 second intervals. The time I announce will be in minutes and seconds remaining. The timer will also furnish you with a running indication of time remaining.

Again, attempt to speak clearly and distinctly. No provision has been made to back up and correct mistakes or misrecognitions, so you must repeat the word or phrase until it is correctly recognized. If you miss a word and don't catch it, I will prompt you for the word you missed: please try again on the word I prompt you with.

Let me re-emphasize here that the shortened time-allowance is meant to represent the urgency of the situation in the Command Post, and that your principal goal is to finish as much of the scenario as possible.

Thanks again for your help.

APPENDIX F

APPENDIX G

QUESTIONNAIRE

****

SUBJECT NUMBER..........

DATE....................

PHASE...................

Please answer the following questions concerning your feelings during the phase of the experiment you have just completed. These questions pertain ONLY to this phase.

Circle the response which best describes your feelings. The scale is:

1 = NOT AT ALL

3 = MODERATELY OR SOMEWHAT

5 = VERY MUCH SO

Responses 2 and 4 are the midpoints between the above.

QUESTION: (Circle one for each question)

During this phase of the experiment, I felt:

	not at all		moderately		very much so
1. nervous	1	2	3	4	5
2. confused	1	2	3	4	5
3. pressured	1	2	3	4	5
4. irritated	1	2	3	4	5
5. anxious	1	2	3	4	5

*****PLEASE! Don't discuss your experience or answers with anyone. ******

LIST OF REFERENCES

1. Levi, L., Emotional Stress, Forvarsmedecin, 1967.

2. Fleishman, E. A., "Introduction", in Alliusi, E. A., and Fleishamn, E. A., Human Performance and Productivity, Vol. III, Earlbaum Associates, 1982.

3. Naval Postgraduate School Report NPS55-82-016, Experiments with Voice Input for Command and Control: Using Voice Input to Operate a Distributed Computer Network, G.K. Poock, April 1980.

4. Lea, W. A., "The Value of Speech Recognition Systems", in Lea, W. A., Trends in Speech Recognition, Prentice Hall, 1980.

5. Illinois University at Urbana Engineering Psychology Labs Report EPL-82-1/ONR-82-1, The Effect of Stimulus-Central Processing- Response Compatability and Resource Competition on Pilot Performance, Sandry, D., and Wickens, C., 1982.

6. Rahe, R. H., "Lifechanges and Near Future Illness Reports", in Levi, L., Emotions - Their Parameters and Measurement, Raven Press, 1975.

7. Lager, C. G., "Experimental Methods and Results of 'Stress' Measurement in Simulated Flight", in Levi,L., Emotional Stress, Forvarsmedecin, 1967.

8. Hicks, J.W., Jr., An Acoustical/Temporal Analysis of Emotional Stress in Speech, Ph.D Dissertation, University of Florida, 1979.

9. Scherer, K.R., "Vocal Indicators of Stress", in Darby, J.K., Speech Evaluation in Psychiatry, Grune and Stratton, 1981.

10. Batchellor, M.P., Investigation of Parameters Affecting Voice Recognition Systems in C3 Systems, M.S. Thesis, Naval Postgraduate School, 1981.

11. Chiles, W.D., "Workload, Task, And Situational Factors as Modifiers of Complex Human Performance", in Alliusi, E.A., and Fleishman, E.A., Human Performance and Productivity, Vol. III, Earlbaum Associates, 1982.

12. Threshold T600 User's Maunual, Threshold Technology, Inc., 1978.

13. Hogan, R., and Hogan, J.C., "Subjective Correlates of Stress and Human Performance", in Alliusi, E.A., and Fleishman, E.A., Human Performance and Productivity, Vol. III, Earlbaum Associates, 1982.

14. Bruning, J.L., and Klintz, B.L., Computational Handbook of Statistics, Scott, Foresman and Company, 1968.

15. Ott, L., An Introduction to Statistical Methods and Data Analysis, Duxbury Press, 1977.

16. Hicks, C.R., Fundamental Concepts in the Design of Experiments, Holt, Rinehart, and Winston, 1973.

17. Conover, W.J., Practical Nonparametric Statistics, Wiley, 1980.

18. Sarason, I.G., de Monchaux, C., and Hunt, T., "Methodological Issues in the Assessment of Life Stress", in Levi, L., Emotions - Their Parameters and Measurement, Raven Press, 1975.

BIBLIOGRAPHY

Apple, W., Streeter, L.A., and Krauss, R.M., "Effects of Pitch and Speech Rate on Personal Attributions", Journal of Personality and Social Psychology, v. 37, pp. 715-729, 1979.

Appley, M.H., and Trumbull, R., eds., Psychological Stress, Appleton Century Crofts, 1967.

Atkinson, J.E., "Inter- and Intra-speaker Variability in Fundamental Voice Frequency, Journal of the Acoustical Society of America, v. 60, pp. 440-445, 1976.

Barrett, A.M., People Under Pressure, Twayne, 1960.

Cooper, C.I., and Marshall, J., eds., White Collar and Psychological Stress, Wiley, 1980.

Cooper, C.I., and Payne, R., eds., Stress at Work, Wiley, 1978.

Eden, G., and Inbar, G.F., "Physiological Model and Analysis of Involuntary Human Voice Tremor", Biological Cybernetics, v. 30, pp. 179-185, 1978.

Friedhoff, A.J., Alpert, M.S., and Kurtsberg, R.L., "An Electro-acoustic Analysis of the Effects of Stress on the Voice", Journal of Neuropsychiatry, v. 5, pp. 266-272, 1964.

Grosser, G.H., Wechsler, H., and Greenblatt, M., eds., The Threat of Impending Disaster, MIT Press, 1964.

Hamilton, V., and Warburton, D.M., eds., Human Stress and Cognition, Wiley, 1979.

Hayre, H.S., "A Preliminary Examination of Rater-determined versus Hayre's Voice-based Psychological Parameters", Applied Acoustics, pp. 63-67, 1980.

Hayre, H.S., and Holland, J.C., "Cross-correlation of Voice and Heart Rate as Stress Measures", Applied Acoustics, pp. 57-62, 1980.

Hicler, N., The Expression of Emotion in the Human Voice, Ph.D Dissertation, University of South Florida, 1978.

Hollien, H., A Case Against Stress Evaluators and Voice Lie Detection, paper presented before the Conference of the Academy for Forensic Applications of the Communications Sciences, Salt Lake City, Utah, December 1, 1979.

Horvath, F., "An Experimental Comparison of the Psychological Stress Evaluator and Galvanic Skin Response in Detection of Deception", Journal of Applied Psychology, v. 64, pp. 323-330, 1979.

Kuroda, I., Fujiwara, O., Okamura, N., and Utsuki, N., "A Method for Determining Pilot Stress Through Analysis of Voice Communication", Aviation, Space, and Environmental Medecine, v. 47, pp. 528-533, 1976.

Levin, H., and Lord, W., "Speech Pitch Frequency as an Emotional State Indicator", IEEE Transactions on Systems, Man, and Cybernetics, March 1975.

Lieberman, P., "Perturbations in Vocal Pitch", Journal of the Acoustical Society of America, v. 33, pp. 597-603, 1961.

Lieberman, P., and Michaels, S.B., "Some Aspects of Fundamental Frequency and Envelope Amplitude as Related to the Emotional Content of Speech", Journal of the Acoustical Society of America, v. 34, pp. 922-927, 1962.

Lippold, C., "Physiological Tremor", Scientific American, v. 224, pp.65-73, 1971.

Ostwald, P., "The Sounds of Emotional Disturbance", Archives of General Psychiatry, v. 5, pp. 587-592, 1961.

Schiffiett, S.G., "Voice Stress Analysis as a measure of Operator Workload", Proceedings of the Human Factor Society Annual Meeting, 1979, also published under U.S. Naval Air Test Center TM 79-3, 1980.

Selye, H., The Physiology and Pathology of Exposure to Stress, Acta Inc. Medical Publishers, 1950.

Shipp, T., and Izdebski, K., "Current Evidence for the Existence of Laryngeal Macrotremor and Microtremor", Journal of Forensic Science, v. 26, pp. 501-506, 1981.

Simonov, P.V., and Frolov, M.V., "Utilization of Human Voice for Estimation of Man's Emotional Stress and State of Attention", Aerospace Medecine, pp. 256-258, 1973.

Williams, C.E., and Stevens, K.N., "Emotions and Speech: Some Acoustical Correlates", Journal of the Acoustical Society of America, v. 52, pp. 1238-1250, 1972.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Technical Information Center Cameron Station Alexandria, Virginia 22314	2
2.	Superintendent ATTN: Library, Code 0142 Naval Postgraduate School Monterey, California 93940	1
3.	Superintendent ATTN: Professor G. Poock, Code 55Pk Naval Postgraduate School Monterey, California 93940	89
4.	Superintendent ATTN: LTC S. Paek, USA, Code 55Ph Naval Postgraduate School Monterey, California 93940	1
5.	Superintendent ATTN: Professor D. Neil, Code 55Ni Naval Postgraduate School Monterey, California 93940	1
6.	Superintendent ATTN: Code 012A Naval Postgraduate School Monterey, California 93940	1
7.	Superintendent ATTN: Code 39 Naval Postgraduate School Monterey, California 93940	1
8.	IBM Communications Products Division ATTN: David W. Davenport PO Box 12195 Dept D48/B632 Research Triangle Park, NC 27709	1

9. SINGER 1
Link Simulation Systems Division
ATTN: Dr. E. Scott Baudhuin
11800 Tech Road
Silver Springs, Maryland 20904

10. Naval Electronics Systems Center 1
ATTN: Frank Deckelman, Code 613
2511 Jefferson Davis Highway
Arlington, Virginia 20363

11. USS Enterprise (CVN 65) 1
Operations Department
ATTN: LT. B. French
FPO San Francisco, California 96635

END

DATE
FILMED

6 - 83

DTIC